

Revisiting Ethical Issues in Research into Social Media Language Use

James McLellan

Faculty of Arts and Social Sciences, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong, Bandar Seri Begawan BE1410, Brunei Darussalam.

ABSTRACT - This editorial discusses workable and unworkable ethical guidelines and shares some narratives about the negotiation of ethical permissions for conducting applied linguistic research using texts from social media platforms. Some of these are from my own experience of over 35 years researching online language use, especially codemixing, alias translanguaging, and practices in language and content subject classes. Among a number of themes emerging from these narratives, the universal applicability of ethical principles and stances needs to be questioned; also, the positioning of the researcher(s): participant- or non-participant-observers, or 'lurkers'. Another issue is the status of online texts, such as those in public online discussion forums, where participants often use pseudonyms: should they be considered as published material and referenced accordingly? Should the pseudonyms appear in cited texts where the content may be contentious? The narratives are presented as vignettes, and it is hoped that we can move beyond the basic dualities of ethical vs. unethical towards a more nuanced analysis.

ARTICLE HISTORY

Received: 7 March 2025

Revised: 11 April 2025

Accepted: 19 April 2025

Published: 24 April 2025

KEYWORDS

Ethics

Language use

Research

Social media

1.0 INTRODUCTION

This paper does not aim to set out any guidelines for ethical conduct of research into social media language use. That would be too prescriptive. In Applied Linguistics we describe, rather than prescribe. Instead, this paper will present examples and evidence and ask questions with the aim of consciousness-raising, in the hope that readers can form their judgments and thereby contribute to ongoing discussions about research ethics when investigating Electronically Mediated Discourse (henceforth EMD).

The scope and focus of the paper need to be made clear at the outset: it is about ethical issues encountered when conducting academic research into social media language use, especially qualitative research analysing social media texts. It does not cover the wider (but equally important) ethical issues connected with participation in and use of social media. Hence the articles in the *Applied Linguistics Review* special issue (Spilioti & Tagg, 2017), entitled "Ethics of Online Research Methods in Applied Linguistics", are of particular relevance. The most thorough and relevant recent study on this topic is that of Strommel and de Rijk (2024): they analysed a sample of 132 journal articles on EMD, finding that only one-third of these included any reference to ethical issues pertaining to data collection and analysis of social media texts. In a useful overview based on the outcomes of a 2-day workshop, Townsend and Wallace (2016, p. 4) note that "given the relatively new and emerging context of social media platforms as research sites, there is as yet no clear ethical framework for researchers entering this field." This can serve as a justification for this present editorial piece.

I will begin by referring to parts of the ethical statement in my own PhD thesis, submitted in 2005, on the topic "Malay-English language alternation in two Brunei Darussalam online discussion forums":

... the postings texts are therefore treated as online publications in respect of text ownership, with the privacy of the text producers respected and maintained through referring to them, wherever this is necessary, only by their chosen pseudonyms. A consequence of this positioning is that there is no possibility of gaining access to the opinions of the text producers themselves, hence ... the analysis consists of text-based analysis of linguistic and discourse product,...

McLellan (2005, p. 30)

This is offered not as a model of good ethical practice, but as a text that you may like to review and comment (critically) on, bearing in mind that social media research and research ethics are fast-moving fields, and that 2005 is a long time ago in relative terms.

Both the forums which served as sources for data texts were public, not password-protected, and hence available to be read by anyone with internet access. As a 'lurker' analysing a random sample of texts posted in the two forums, I had no means of knowing the identity of the text producers, as they invariably used pseudonyms. At the time of analysing the texts and writing up the findings, I felt that this was sufficient to safeguard anonymity and confidentiality. If such a study were replicated in 2025 I would most likely blank out the pseudonyms, and any other identifying references (such as "xxxx, where I live....").

Though perhaps not very useful, it is possible to go back even further in time and investigate the development of ethical research practices. The most likely avenue to explore here is perhaps that of transfer from the fields of medicine and other experimental hard science research, where there is a strong desire to get away from horrendously unethical research practices such as those of the 1932-1972 Tuskegee Experiment in the USA (<https://www.cdc.gov/tuskegee/about/index.html>).

In social media language research, perhaps, we sail in less troubled waters, as there is less likelihood of physical harm being done to participants through studies of the languages they use on social media platforms. But there are nonetheless issues to navigate which are central to research ethics, including public and private texts, informed consent, confidentiality and anonymity.

This paper covers some of these issues relating to both research by undergraduate and graduate students, and to research by academic staff. For both categories, there are many commonalities: nowadays all universities have procedures for reviewing research studies to ensure that they comply with ethical guidelines (see examples below).

The next distinction that must be drawn is between public and private social media texts, in line with the assumption that public texts, posted online in the knowledge and hope of their being read, are usable as data, whereas private texts, as listed in Table 1, require the informed consent of text producers. Table 1 distinguishes between these:

Table 1

Public vs private social media sites

Public	Private
Blogs	Facebook groups
Online discussion forums (e.g. Reddit)	Facebook Messenger
Instagram	Whatsapp, Viber, WeChat chats and groups
Facebook status updates (walls)	Text messages (SMS)
YouTube (videos, & comments below videos)	
X (Twitter)	
TikTok	

It must be noted here that the private/public distinction can be problematic (Giazoglou, 2016). There is a further ongoing debate about the ethics of direct quotations from public social media texts versus paraphrasing: direct quotations, even with nicknames or pseudonyms blanked out, can lead to the identity of the text producer being revealed (Strommel, 2024). This is a sentence from a Brunei discussion forum (Reddit) which I recently used while analysing translanguaging (language alternation),

[1]

"Old habits die hard nya urang, tapi *it gets worse* kalau dorang ani merasa inda diterima lagi".
3P but if 3P DEM AV-feel NEG VPass-receive again

(free translation: 'old habits die hard for these people, but it gets worse if they feel unwelcomed again' - *source suppressed*)

A search for this sentence on 30th August 2021 in Google led, via the first hit, directly back to the page of the forum in which it was posted, showing the Reddit nickname of the message poster. My interest is in the three switching points: which components are in (Brunei) Malay and which are in English, and in the use of the clichéd idiom "Old habits die hard". The topic of the discussion thread from which this is taken, discrimination by society against ex-prisoners, is not part of my analysis. So, does that make it acceptable to cite it directly? Should we go against our natural academic preference for direct quotation to ensure an accurate representation of others' words, and opt for paraphrasing instead?

2.0 RESEARCH INTO SOCIAL MEDIA LANGUAGE IN ASIAN CONTEXTS: SAME ETHICAL STANDARDS AND EXPECTATIONS?

We need to question the universality of ethical principles relating to social media research: are they global, or locally variable? Here we can compare research ethics to plagiarism, a topic which scholars such as Pennycook (1996), among others, have shown to be variable across languages and academic cultures. Both topics, social media research ethics and plagiarism, converge in the area of text ownership and the right to cite.

Classroom interaction is another comparable field: it is central to Applied Linguistics, which is traditionally defined as applications of Linguistics to Language Learning and Teaching. Ethical issues in research into classroom interaction also demonstrate variability across lingua-cultures. In the academic culture of New Zealand, the informed consent of teachers, students and all parents is a prerequisite for undertaking any study using audio- or video-recorded data. In Vietnam, by contrast, teachers are held in particular reverence and high esteem and are deemed *in loco parentis* and thus able to give consent on behalf of the students in their charge (Huong Thi Nguyen, 2021, p. 30).

A recent study by Bravo-Sotelo (2024) on translanguaging practices in the Philippines focused on Year 8 Mathematics classes conducted online during the COVID-19 pandemic. For this, the researcher took care to observe ethical principles and practices, as explained here:

After accomplishing and submitting the research ethics form to the faculty committee, the clearance form was approved and obtained. I then sought permission from the Philippine Department of Education through the school principal and provided

appropriate consent forms to the research participants. For the teachers, they immediately sent back the signed forms after the orientation. For the students, since they were below 18 years old, the teachers gave each one a consent form to be signed by their parents. Before the teachers gave out the forms, I was given the chance to introduce myself in each class and to explain in Filipino the details of the research, such as the procedures and implications. Likewise, I mentioned that the participation was fully voluntary and they could withdraw from the research at any point without any consequence. All student-participants sent back the consent forms.

Bravo-Sotelo, (2024, p. 37)

With the move to online learning and teaching made necessary by the recent Covid19 pandemic, this has become a social media issue also: permissions (from whom?) are required to use online interactional data from learning and teaching contexts for research purposes.

2.1 English-Only?

In the highly multilingual context of Southeast Asia, we should be conscious that not all social media, and not all research, operates through the medium of English (Danet & Herring, 2007). Having a major research interest in language alternation (code-switching, translanguaging) in online EMD contexts, I have to be aware of the danger of Anglocentrism. Hence for this editorial I have tried to search for ethical guidelines for social media language research in Malay/Bahasa Indonesia. While there are plentiful studies about the broader area of ethics relating to social media language use and participation, I have thus far found nothing in Malay specifically on the narrower topic of this paper, ethical issues pertaining to research into social media language use, despite having done searches using strings such as "*etika kajian bahasa media sosial*". I would therefore like to invite others to share their insights based on guidelines in Thai, Malay, Chinese, Vietnamese, Filipino, Hindi, Tamil, Urdu, and other major Asian languages which are found in social media and used for disseminating research findings.

3.0 RESEARCH BY UNDERGRADUATE AND GRADUATE STUDENTS

All universities now require undergraduates, especially those undertaking final-year research projects, as well as graduate research students on Masters- and PhD-level programmes, to apply for ethical clearance. The procedures for obtaining this may be complex: Figure 1 illustrates an example from our FASS faculty's website at Universiti Brunei Darussalam.

Figure 1

Ethics approval process, Faculty of Arts and Social Science, Universiti Brunei Darussalam

All research involving humans or animals must undergo ethics approval.
Research conducted by FASS PhD and Masters students should be submitted to the FASS Ethics Committee.

1.1.1 Ethical clearance procedures:

- Step 1: Students first consult with their supervisor regarding the proposed proposal to check if it requires any ethical clearance.
- Step 2: Having read the proposal the supervisor should decide if there are any ethical concerns in the proposal that need to be forwarded to the FASS ethics committee for clearance.
- Step 3: If the supervisor thinks that the proposal needs to be forwarded to the ethics committee, he/she should suggest that the student download the ethics clearance form from the FASS website, fill it up and send to the ethics committee.
- Step 4: The student should fill up the hardcopy ethics clearance form, attach the relevant documents, and send them to FASS General Office (Room 2.93). All soft copy of the forms should be sent via email (office.fass@ubd.edu.bn).
- Step 5: The committee will review the proposal and will inform the student about the outcome as soon as possible via email.

Source: <https://fass.ubd.edu.bn/links/ethics.html>

This outlines only the first stage: following provisional approval, if granted, by the faculty-level committee, the application must be sent up to the university-level ethics committee. This of course can be time-consuming. The text in Figure 1 above refers repeatedly to the role of the supervisor in ensuring compliance in student research. Hence supervisors are charged with ensuring that their research students are made aware of and comply with all ethics procedures and requirements.

King Mongkut's University of Technology, Thonburi, Thailand (KMUTT), had this diagram in 2021 to explain its ethical expectations and practices with reference to all research, not just social media language research (see Figure 2):

Figure 2

Research Ethical Expectations and Practices at King Mongkut's University, Thailand



Source: <https://ethics.kmutt.ac.th/wp-content/uploads/2021/05/Principle-of-Human-Research-Ethics-ENG.pdf>

As an example of research into private EMD, language choices in Whatsapp chats and groups, 'Aqilah Aziz (2019, p. 29, cited with permission) summarizes the research ethics aspect of her study by explaining,

Participants were asked beforehand for permission to use their chat logs for the purpose of data research as it concerned their privacy... The time frame of the chat logs... were from months prior to the start of the research, so participants did not have the chance to adjust their conversations to the needs of the research. However, they did have the option to remove any sensitive issues which they might not want the researcher to know... Data such as names and places are made anonymous to ensure their privacy.

(Aqilah Aziz, 2019, p.29)

She only used the Whatsapp paired chats and groups in which all participants had given their consent, and in her data transcripts, included as appendices, names were replaced by the designations F1, F2, F3 and M1, M2, M3 for female and male participants respectively, as shown in this brief extract from a group chat:

- M3: Abis bat kah?
 Finish ABBR-battery DM - question-tag
 ('Is the battery dead?')
- M3: Or rusak
 Broken
 ('or is it broken?')
- F5: Mcm rosak
 ABBR-like broken
 ('It's sort of broken')
- F5: My Adik ckp
 younger sibling ABBR-said
 ('said my younger sibling')
- F6: [photo] Dunno if habis batt or rosak
 finish ABBR-battery broken
 ('don't know if the battery is dead or broken')
- F2: I think mine tinggal di uk pasal rosak sudah
 left in because broken already
 ('I think mine's left in the UK because it was broken already')

As well as anonymizing names, a photo of the participant ("F6", in line 5) was deleted because this could be an identifier. Another PhD researcher (not named, as permission cannot be obtained) was given the informed consent of students taking a TESL programme in a Malaysian university to use their texts posted in a Facebook group as research data.

They were informed beforehand that they are invited to participate in the study and they will remain anonymous. I also briefed them that they are free to agree or disagree to participate in the study or in the Facebook group. ... Hence, to serve the purpose of the study, the walls of 50 Facebook users who are (university) TESL undergraduates were analysed using document/content analysis.

Source: PhD thesis, author's name suppressed

Both this researcher and 'Aqilah Aziz were thus able to conduct their research into language choice and use, including language mixing, in social media platforms deemed private in Table 1.

For online questionnaire surveys, including those sent out electronically via Google Docs, Qaltrics, or SurveyMonkey, participants are deemed to have given their informed consent by agreeing to fill out the survey form and submit their answers anonymously: they are under no coercion so to do.

In such instances, online technologies are used as mediums for research data collection: EMD is not necessarily the research topic or focus. For both student and academic staff research, surveys followed by interviews with a smaller number of participants are regularly used methodology choices. Whilst the survey results can remain anonymous, the information given or elicited in an interview, then cited and discussed, may cause the participant to be identifiable, unless measures are taken to avoid this, for example by deleting information divulged by participants about their home and family background.

4.0 CONCLUSION

There are other stakeholders and concerned parties in social media language research ethics. Firstly, the researcher: ethical dilemmas cut both ways, researchers have rights also. If the restrictions are too stringent, then they cannot do their job as researchers. For example, how can the researcher's rights be ensured, and the moral and ethical dilemma resolved if a research interview reveals illegal acts such as drug peddling? However, in the Malay-Muslim context of Brunei, Malaysia and Indonesia, the obligations and expectations of researchers are paramount. Noor Azam Haji-Othman and Azmi Mohamed (2021, p. 125) note that

In Bruneian-Malay terms, it is the researcher who is *behajat* – the one who needs something, in this case, precious data. Within the Malay-Muslim context of Brunei, *awar-galat* (proprietary), *berbudi-bahasa* (being respectful), and *bersopan-santun* (having good manners) are of utmost importance, constituting a large part of the Malay cultural system of politeness in human interaction, including the specific context of research interviews.

This begs the question of whether and how researchers into social media language use can meet their culturally determined obligations when not interacting directly, face-to-face with their participants or consultants.

Secondly, publishers of journals and books will demand and expect that research on social media language use conforms to ethical norms, as they might otherwise be legally liable. Publishers and journal editors thus expect authors of articles submitted to take responsibility for ensuring ethics compliance, and any questioning of this on the part of peer reviewers may cause the submitted article to be rejected.

In light of the perspectives and examples above, there are at least two possible future directions for this discussion to take. Firstly, further exploring and questioning the core elements of research ethics as they may apply to social media:

- whether informed consent is needed or not, and how this is requested and granted
- what counts as private vs what counts as public for texts used as data for linguistic and discursual analysis
- measures which can be taken by researchers to maintain the anonymity and confidentiality of their participants.
- whose informed consent is required for research into online classroom interaction

Finally, there is the seriously under-researched topic of research ethics across languages and academic cultures. In worst-case scenarios, ethics regulations and guidelines drawn up and imposed by younger and newer universities have been found to be themselves plagiarized from the websites of established universities with higher ranking, leading to the neocolonial imposition of ethical research standards which may not be well aligned with the local academic culture. This short exploration concludes where it started, with a disclaimer about being prescriptive. I echo Townsend and Wallace (2016, p. 15) who conclude: "This framework, or any framework that supersedes it, cannot be prescriptive".

ACKNOWLEDGEMENTS

My sincere gratitude goes to those who agreed to their data and ethical statements being cited in this editorial. This study was not supported by any grants from funding bodies in the public, private, or not-for-profit sectors.

CONFLICT OF INTEREST

There is no conflict of interest in this sole-authored editorial.

AUTHOR CONTRIBUTIONS

James McLellan (Conceptualisation, Investigation, Resources, Writing)

REFERENCES

- 'Aqilah Aziz (2019). *Language choice and code switching in Brunei in synchronous electronically mediated communication*. PhD thesis, Universiti Brunei Darussalam.
- Barnard, R., & Yi Wang. (Eds, 2021). *Research ethics in Second Language Education: Universal principles, local practices*. Routledge.
- Bravo-Sotelo, K.P. (2024). Investigating the translanguaging practices of mathematics teachers and students in online classes in the Philippines. Unpublished PhD thesis, FASS, Universiti Brunei Darussalam.
- Danet, B., & Herring, S. (2007). *The Multilingual Internet: Language, culture and communication online*. Oxford University Press.
- Giaxoglou, K. (2016). Reflections on internet research ethics from language-focused research on web-based mourning: revisiting the private/public distinction as a language ideology of differentiation. *Applied Linguistics Review*, 8(2–3), 229–250. <https://doi.org/10.1515/applirev-2016-1037>
- Huong Thi Nguyen (2021). Oral corrective feedback: Ethical issues in researching Vietnamese lecturers' beliefs and practices. In R. Barnard & Yi Wang (Eds.). *Research ethics in Second Language Education: Universal principles, local practices* (pp. 28-44). Routledge.
- McLellan, J. (2005). *Malay-English language alternation in two Brunei Darussalam online discussion forums*. Unpublished PhD thesis, Curtin University of Technology.
- Noor Azam Haji-Othman & Azmi Mohamad (2021). Research ethics from a Malay-Muslim perspective. In R. Barnard & Yi Wang (Eds.). *Research ethics in Second Language Education: Universal principles, local practices* (pp. 117-127). Routledge.
- Pennycook, A. (1996). Borrowing Others' Words: Text, Ownership, Memory, and Plagiarism. *TESOL Quarterly*, 30(2), 201. <https://doi.org/10.2307/3588141>
- Sormanen, N., & Lauk, E. (2016). Issues of Ethics and Methods in Studying Social Media. *Media and Communication*, 4(4), 63–65. <https://doi.org/10.17645/mac.v4i4.793>
- Spilioti, T., & Tagg, C. (2016). The Ethics of Online Research Methods in Applied Linguistics: Challenges, opportunities, and directions in ethical decision-making. *Applied Linguistics Review*, 8(2–3), 163–167. <https://doi.org/10.1515/applirev-2016-1033>
- Stommel, W. (2024). Social media research and ethics: What does the user want? Accessed 6 March 2025 via <https://researchmethodscommunity.sagepub.com/blog/social-media-research-ethics-user->
- Stommel, W., & De Rijk, L. (2021). Ethical approval: none sought. How discourse analysts report ethical issues around publicly available online data. *Research Ethics*. <https://doi.org/10.1177/1747016120988767>
- Townsend, L., & Wallace, C. (2016). *Social media research: A guide to ethics*. Aberdeen, Scotland: University of Aberdeen. Accessed 6 March 2024 via https://www.gla.ac.uk/media/Media_487729_smxx.pdf
- Zimmer, M. (2020). "But the data is already public": on the ethics of research in Facebook. *The Ethics of Information Technologies*, 229–241. <https://doi.org/10.4324/9781003075011-17>